The effect of ethnic diversity on income – an empirical investigation using survey data from a post-conflict environment

Adnan Efendic and Geoff Pugh

Abstract

This empirical study is based on nationally representative cross-sectional survey data gathered to investigate the effect of ethnic diversity on personal and family incomes in Bosnia and Herzegovina (BiH), a post-conflict society. The 1992–1995 conflict was harmful for ethnic diversity. Yet, two decades later, where it still exists, ethnic diversity gives rise to positive economic consequences. After controlling for other influences, the authors find lower probabilities of respondents in ethnically diverse neighbourhoods being in the lowest income categories but higher probabilities of being in medium and higher income categories. The largest effects are the reduced probabilities of respondents reporting no income, and are thus particularly relevant to poverty reduction. At the limit, their estimates imply an income gap of more than 20 per cent between a counter-factual completely heterogenous environment and a counter-factual completely homogenous environment. Policy makers in this post-conflict country, and in similar environments elsewhere, should take into consideration the economic costs of policies supporting ethnic homogeneity over diversity.

Authors
Adnan Efendic, School of Economics and Business, University of Sarajevo, Sarajevo, Bosnia and Herzegovina, adnan.efendic@efsa.unsa.ba
Geoff Pugh, Staffordshire University, Stoke-on-Trent, Great Britain

1   Introduction

Ethnic diversity is an ongoing research topic in economics, investigated both at macro and micro-economic levels, and is a topical issue among policy makers. Published findings on the economic consequences of ethnic diversity remain mixed, which suggests the need for further empirical research into the potential welfare outcomes of ethnic diversity. This study contributes in the specific context of a post-conflict environment.

In a related study, Efendic et al. (2015) find that ethnic diversity in the post-conflict environment of Bosnia and Herzegovina (BiH) is beneficial for young businesses. Young companies (entrepreneurs) in local areas that are ethnically diverse have systematically higher growth aspirations in comparison to those in ethnically homogenous areas. These results are important when seen in the context of the cross-country economic literature, as the latter suggests that ethnic heterogeneity may be associated with negative economic outcomes (e.g. Easterly and Levine 1997; Collier 1998; Easterly 2001; Patsiurko et al. 2012). Motivated by this research and a growing body of literature that investigates the effect of ethnic diversity in economics, we conducted a new survey of the household sector in BiH to investigate whether individual and household economic performance is associated with ethnic diversity in the local area. We investigate a country that has been ethnically diverse for centuries (Malcolm 1996). Yet, two decades ago, this ethnically heterogenous composition was changed as a result of the Bosnian war (1992–1995), which involved ethnic cleansing and related mass fatalities (Olzak 2011). As yet, the economic outcomes for individuals and families of consequently reduced ethnic diversity in BiH remain unknown (Cvitkovic 2017: 45).

Our study contributes to the literature by using unique survey data to analyse the effect of ethnic diversity on individual and household incomes in a post-conflict society. Previously, moreover, ethnic diversity has been investigated mainly in the framework of ethnic heterogeneities caused by immigration and inflow of different cultures and traditions into formerly homogenous areas.[1] This research treats the economic consequences of change in the opposite direction – namely, a shift towards homogenisation in much of BiH – which is something new, being enforced largely exogenously by violence (Malcolm 1996). Thus, the findings from this study have implications for societies facing similar changes. Moreover, while there are already a good number of studies at the macro (country) level (e.g. Easterly and Levine 1997; Collier 1998; Alesina and La Ferrara 2005; Montalvo and Reynal-Querol 2005; Patsiurko et al. 2012; Gören 2014), this is the first research, to our knowledge, to investigate the effect of ethnic diversity at the individual and household levels. An advantage of this approach is the much greater variation in ethnic diversity than at aggregate levels. While it is not our intention to explain the contrasting findings of macro and micro studies (discussed below), we conjecture

---

[1] In the literature reviewed above, many authors (e.g. Alesina and La Ferrara 2005) distinguish ethnic from linguistic and religious groups in their research and analyse these differences separately. In the context of BiH, however, there are no "real" language differences inside the country between different ethnicities. In contrast, the dominant ethnicities in BiH (Bosniaks, Serbs, and Croats) largely correspond to religious differentiations (Muslims, Orthodox, and Catholics respectively).

that any relationship between ethnic diversity and economic performance is less likely to suffer from omitted variables bias at the micro than at the macro level, which dominates existing research.[2]

The paper is organized as follows. Section 2 provides background on the post-conflict context of BiH, in which we explain why this country is relevant for such an investigation. Literature review in Section 3 suggests that ethnic diversity affects economic performance in both negative and positive ways. Section 4 explains the survey design and the data collected for this study, while Section 5 explains how theoretical concepts are matched to variables derived from survey questions for estimation. In Section 6 we report and discuss the empirical findings. Section 7 reports extensive robustness checks based on alternative model specifications and methods of estimation, including an instrumental variables approach to investigate the potential endogeneity of our variable of interest. The interpretable quantitative results of this investigation are reported in Section 8, while Section 9 concludes by summarizing key empirical findings and considering policy implications.

2 The post-conflict context of Bosnia and Herzegovina

To investigate the research question of interest – whether ethnic diversity is associated with individual and household economic performance – we focus on Bosnia and Herzegovina, a post-conflict country, which is a particularly appropriate context for such a study.[3] Throughout its one thousand years long history, BiH had been recognized as a multicultural environment, mixing Eastern and Western cultural and religious influences (Malcolm 1996). Even today, BiH is a multicultural country with a complex institutional and ethnic structure, and great variations within the country. According to the latest census data (2013), three ethnic groups dominate: around 50% are Bosniaks; 31% Serbs; and 15% Croats. It is worth mentioning that the ethnic identities in BiH were mainly built on religious identities: Bosnian Muslims most often declare themselves as Bosniaks; Bosnian Catholics as Croats, and Bosnian Orthodox as Serbs. Institutionally, BiH is composed of two entities and one district (meso-structure), where Bosniaks and Croats dominate in the Federation of BiH (FBiH) entity, and Serbs in the Republika Srpska (RSBiH) entity, while the District Brcko of BiH (DBBiH) mixes all three ethnicities with no group dominant.

_____

[2] For example, at the macro level high rates of economic growth favour inward migration potentially causing ethnic diversity. At the micro level of our study – the "neighbourhood" – differences in economic growth might impact travel to work patterns rather than cause people to move home (i.e. to "migrate" to another neighbourhood). In this case, the joint determination of individual economic performance and ethnic diversity by economic growth is more likely at higher than at lower levels of aggregation.

[3] See Efendic et al. (2017) for wider discussion of migration and ethnic diversity in South East Europe within a theoretical framework informed by social capital.

When BiH was part of former Yugoslavia (1945–1992) it was a republic with 4.1 million citizens and was particularly well-known for its multi-ethnic, multi-cultural, multi-religious environment, as well as for being the republic with the highest level of ethnic tolerance (Hodson et al. 1994; Dyrstad 2012). The Bosnian war (1992–1995),[4] which followed the dissolution of former Yugoslavia, caused a structural break in the demographic and ethnic composition of the country; namely, a change from ethnic diversity to ethnic homogeneity in some areas. This was a consequence of large migration movements induced by the war, including outflow of population from BiH (estimated at around 1.2 million over the period 1992–1995) in addition to around 1.0 million internally displaced. After the Dayton Peace Agreement was signed in 1995, people started to return; in the following five years it is estimated that around 1.0 million returned to BiH (FMDPR 2011). Accordingly, close to 50% of the pre-war BiH population migrated during the war period. The nationally representative survey data used for this study (explained in detail later) provide a consistent massage: 11% of respondents left their homes during the war period but returned shortly afterwards; 22% returned but a few years after the conflict; 12.5% never returned to their pre-war homes; and the remaining 52% in our sample did not move. These 12.5% who did not return to their pre-war homes changed the diversity picture in both places, both in their home municipalities and in their new municipalities of residence. In our sample, 93% of this group self-identify as belonging to the majority ethnicity in the areas where they live today, most likely having moved from places where they would have belonged to a minority. We find that 54% of municipalities host such war migrants, with an almost identical distribution between the two entities (54% of municipalities in RSBiH and 53% in FBiH). However, if we examine their distribution at municipal levels, the highest percentages of these war migrants in both entities are recorded in the two biggest municipalities: namely, Sarajevo and Tuzla in FBiH; and Banja Luka and Bijeljina in RSBiH. Finally, comparison of data from the pre-war and post-war censuses reveals 20 from 141 municipalities that not only experienced diversity change but also a shift in the dominant ethnic group: in RSBiH, 12 municipalities changed from a Bosniak to a Serb majority; while in FBiH changes took place towards either a Bosniak majority in 4 municipalities or a Croat majority in another 4.

The war period also induced a structural change in attitudes such that a multi-ethnic BiH went from ethnically quite tolerant to quite intolerant in just a few years (Dyrstad 2012). Two decades after the Bosnian war, the country still remained segregated along ethnic lines, where the three main ethnic groups have substantial autonomy and control over their own ethno-territorial units (Bieber 2010). Yet, in spite of all these changes, there are still a few regions within the country in which ethnic diversity is still preserved and did not change a lot. (In Section 4 below, we present further detailed evidence on changes in ethnic diversity between the censuses of 1991 and 2013 across 127 municipalities.) Accordingly, variations in ethnic diversity between different areas – in particular, at the level of micro-units – constitute a fertile terrain for identifying the potential influence of ethnic homogeneity/diversity on the economic performance of individuals and households living in these areas.

_____

[4] For reasons of space, we do not discuss a deeper classification of this war.

3   The literature on ethnic diversity

The association between different forms of ethnic diversity and economic performance has been the focus of considerable economic research over the last two decades, generally finding that ethnicity does matter. However, the empirical research supports opposing hypotheses regarding ethnic diversity and economic performance, suggesting both positive and negative effects on outcomes as well as more or less strong and/or significant influences (Constant et al. 2009; Olzak 2011). This may reflect attempts to identify economic consequences of ethnic diversity at different levels of aggregation: macro (mainly focused on between-country variations in growth); and micro (region, sector or firm-level variations in economic performance). However, the full heterogeneity of contexts and dimensions of diversity have yet to be researched. Table 1 gives an overview of the literature to be discussed in this section.

A common theoretical proposition of the diversity literature is that – ceteris paribus – greater ethnic diversity increases the probability of ethnic tensions and conflicts (Blimes 2006) which, in turn, have a negative impact on economic incentives and economic performance (Osborne 2000). In this case, ethnic diversities, fractionalisation, conflicts and prejudices can override economic incentives, leading to poor economic choices, policies, outcomes, and political instability. In general, therefore, ethnically polarized societies are more likely to select suboptimal economic policies, which reduces economic prosperity (Easterly and Levine 1997). Accordingly, ethnic diversity is usually associated with poorer economic performance and lower national economic growth (Collier 1998; Alesina and La Ferrara 2005; Goren 2014).

A number of empirical studies report a negative effect of ethnic diversity on economic outcomes. Easterly and Levine (1997) focus on ethno-linguistic diversity at the national level and find that ethnic diversity is associated with slow economic growth in Africa. Moreover, the effect of ethnic diversity is negative not only in its direct effect on economic growth, but ethnic diversity partly explains variations in economically relevant indirect indicators such as schooling, political stability, financial systems, foreign exchange markets, government consumption and infrastructure. On this argument, ethnic diversity can exert indirect effects by influencing the operation of channels or policies that affect long-run growth rates. In line with this, Goren (2014) identifies a direct negative effect of ethnic diversity on economic growth in a global sample as well as a number of indirect transmission channels through which diversity may affect growth – namely, schooling, political instability, market distortions, trade openness and the fertility rate. In this vein, Escaleras and Register (2011) find that ethnic tensions negatively affect the formation of social infrastructure (e.g., public utilities, education, and health care), thereby imposing an unnecessary burden on growth and development. Similarly, Alesina et al. (1999) investigate a sample of US cities and find that greater ethnic diversity in US local jurisdictions is associated with higher spending and higher deficits/debt per capita, but still with lower provision of the core public goods like education and roads.

Collier (1998) in a cross-sectional study of the effect of ethnic diversity on economic growth finds that maximally diverse societies grow more slowly than do homogenous societies.

Table 1: Empirical studies on ethnic diversity and economic performance

Author(s)	Sample – level of analysis	Type of data	Explained variable	Diversity effect reported
Macro-level studies				
Alesina and La Ferrara 2005	Global, country and US county - macro and micro	Quantitative	GDP growth Population growth	Negative for GDP growth at macro level, while interaction with income positive at lower level
Easterly and Levine 1997	Global, country - macro	Quantitative	GDP growth	Negative direct effect, but indirect effect more robust
Gören 2014	Global, country - macro	Quantitative	GDP growth	Negative direct effect, while indirect effect differs
Escaleras and Register 2011	Global, country - macro	Quantitative	Social infrastructure	Negative effect of ethnic tensions
Collier 1998	Global, country - macro	Quantitative	GDP growth	Negative effect reported, but not in democratic societies
Easterly 2001	Global, country - macro	Quantitative	GDP growth	Negative direct effect, but it interacts with institutions
Patsiurko et al. 2012	OECD, country, panel - macro	Quantitative	GDP growth	Negative direct effect, while indirect effects discussed
Montalvo and Reynal-Querol 2005	Global, country, panel - macro	Quantitative	GDP growth	Negative direct and indirect effects reported
Micro-level studies				
Alesina et al. 1999	US cities, metropolitan and county - micro	Quantitative	Public services	Negative effects of ethnic fragmentation
Smallbone et al. 2010	London, four businesses - micro	Qualitative	Competitiveness	Positive effect, but more evidence needed to support it
Ottaviano and Peri 2006	US, cities - micro	Quantitative	Average wage	Positive effect for wages of US-born workers
Bellini et al. 2012	EU, regions - micro	Quantitative	GPD level	Positive effect for productivity of EU regions
Efendic et al. 2015	Post-conflict country, BiH - micro	Quantitative	Growth aspirations of entrepreneurs	Positive effect for growth aspirations of young companies
Source: The authors.				

However, he also finds that diversity is damaging to growth primarily in the context of limited political rights, while this effect is not clearly identified in democratic societies. Easterly (2001) also identifies the negative effect of diversity on economic growth, but the authors report that it is not an isolated effect and might be mitigated by good institutions. Yet Patsiurko et al. (2012) report a negative association between ethnic fractionalization and economic growth for OECD economies. In addition, the authors identify the greater importance of "ethnic fractionalization" in comparison to other forms of fractionalizations, such as religious and linguistic. Similarly, Montalvo and Reynal-Querol (2005) in their cross-country research likewise find that ethnic polarization has a negative effect on economic development. These authors argue that ethnic polarization reduces investment, increases government consumption and entails a higher probability of civil conflict, which ultimately reduce economic development. Although the above discussed literature identifies a negative effect of ethnic diversity on economic performance, the perspective of these studies is primarily macro-economic and often focused on indirect influences on economic outcomes (e.g. through its effect on government efficiency and provision of public goods and services).

A different perspective in the literature is that most developed countries, regions and cities today are ethnically diverse. Proponents of this approach explain that a diverse ethnic mix may bring various abilities, different experiences, a variety of cultures and traditions, a spectrum of religious beliefs and practices, and multidimensional ways of thinking, which together may lead the whole society towards greater innovation, creativity and economic performance. In this case, ethnic diversity might be considered as an important asset for human development and welfare (Alesina and La Ferrara 2005; Bellini 2012). Ethnic diversity might have positive consequences not only at the national level but, in particular, on the economic success of regions and cities (Jacobs 1961; Gertler et al. 2002; Smallbone et al. 2010) as well as on the productivity of individuals (Ottaviano and Peri 2006; Bellini et al. 2012) and, accordingly, individual well-being (Akerlof and Kranton 2010).

Micro-level studies tend to be supportive of themes introduced by Jacobs (1961) who identified ethnic diversity as the key factor of success of a city and as an engine of urban development. The positive potential of ethnic diversity is also suggested by experiments in the business literature that investigate differences in creativity between homogenous and diverse (small) groups of people and report better performance of diverse groups (e.g., McLeod and Lobel 1992; McLeod et al. 1996). Alesina and La Ferrara (2005) argue that more diverse groups with limited abilities can perform better than can more homogenous groups of high-ability problem solvers. Hence, individuals involved in more diverse groups, networks and environments can find better solutions to difficult problems, and so become economically more productive. The same authors develop a theoretical framework in which the skills of individuals from different ethnic groups are complementary in the process of production, which in turn increases productivity. Based on this framework, Alesina and La Ferrara (2005: 763) "verify the consistency of this theory using repeated cross-sectional data on countries and localities in the United States and we find that, while ceteris paribus increases in ethnic diversity are associated with lower growth rates, the interaction between diversity and the income level of the

community under study is positive. This suggests that ethnic diversity can be beneficial (or at least less detrimental) at higher levels of development."

Less differentiated results are reported by Ottaviano and Peri (2006) who aggregate census data across groups of US born workers and find that ethnic diversity is associated with higher wages of the resident population. Bellini et al. (2012) conduct similar research focusing on European regions in twelve European Union countries and find consistent results – namely, ethnic diversity is positively correlated with productivity, where causation goes from diversity to productivity. Finally, at the firm level, Smallbone et al. (2010) provide small-sample, qualitative evidence that ethnic diversity is a source of creativity and innovation among first- and second-generation immigrant entrepreneurs in London; and Efendic et al. (2015) find that ethnic diversity has a positive effect on the growth aspirations of young businesses in BiH.

The previous literature suggests economic mechanisms whereby ethnic diversity promotes creativity and innovation at the micro level, thereby inducing greater productivity of individuals and, hence, business development, job creation and economic prosperity (Smallbone et al. 2010). On the one hand, ethnic diversity provides an environment enlarging the scope for individuals to generate new combinations of ideas and resources, thereby increasing creativity and innovation by individuals, groups and companies. On the other, ethnic diversity makes a location more attractive for a wider range of individuals and entrepreneurs, who are thereby incentivized to base the whole spectrum of their creative identities (artistic, cultural, technological, and generally economic) in such areas (Gertler et al. 2002).

Religion has been identified as an important factor promoting greater openness, innovation, creativity (Korman 2015; Cinnirella and Streb 2017) and economic growth (Barro and McCleary 2003). Because religious and ethnic diversities in BiH intersect, the main religious festivals are – to a large extent – celebrated as cultural events across all communities, which can thus provide micro level events supporting the economic performance of individuals. Donia (2006: 4) recognizes that this tradition and practice of common life has been preserved at most times and in widely varying circumstances, even during the last war. He explains, using Sarajevo as an example: "Sarajevans have long used the concept of neighbourliness to express their respect for those of different faiths and nationalities, manifest in the practices of mutual visitations and well-wishing on holidays as well as everyday cordial relations. Common life is neighbourliness writ large. It embodies those values, experiences, institutions, and aspirations shared by Sarajevans since the city's founding."

In the language of social capital theory, in ethnically diverse neighbourhoods, such multi-cultural sharing of religious festivals may promote "weak ties", which are an acknowledged source of knowledge exchange, innovation and economic betterment (Granovetter 1973; Rebmann et al. 2017). Recent research in the context of BiH provides supportive evidence for these linkages. Efendic (2018) reports that ethnically more diverse networks of individuals in BiH have a positive effect on social capital outcomes at local level, while recent research among entrepreneurs in BiH implemented as part of the INFORM (2017) project confirms that

entrepreneurial networking is widely used but blind to ethnic and religious differences.[5] This evidence that ethnic diversity favours the accumulation of social capital, and that networking across ethnic differences is an important business activity for entrepreneurs, suggests that ethnic heterogeneity at the neighborhood level areas promotes better informal institutions based on diverse informal networking – one of the defining features of social capital – which, in turn, may facilitate economic betterment.

A link between diversity and economic betterment is strongly supported by recent evidence from economic history. Mokyr (2016) argues that the origins of the late 18th Century innovations in Europe that triggered the Industrial Revolution and subsequent economic progress are to be found in the emergence of a "culture of growth" in early modern Europe (c.1500–c.1700). Mokyr (2016: 40–41 and 53) draws on social capital theory to establish the influence of the quantity and quality of interconnections (especially weak ties) between people in promoting innovation (technical progress) and explicitly identifies diversity and pluralist values as essential features of a culture capable of promoting economic dynamism.

"One important cultural value is pluralism: whether to tolerate incompatible values and beliefs, and whether to give new cultural elements – no matter how outrageous they sound – a fair chance to compete in the market place for ideas and values is itself a value that needs to be accepted. A belief in cultural (including religious) tolerance and free speech and thought, and the institutions it implies … can be of great economic value when it is relatively rare: it allows an economy to attract and absorb religious and political refugees, who tend to be creative and well networked. The willingness of the Netherlands, and later Britain and the United States to tolerate Jews and dissenting Christians contributed a great deal to their economies, especially in high-skilled manufacturing and financial services."

Mokyr (2016: 53) footnotes this passage by quoting from Cipolla (1972: 52). "Throughout the centuries the countries in which intolerance and fanaticism prevailed lost to more tolerant countries the most precious of all possible forms of wealth: good human brains … Inflow of good brains and receptiveness to new ideas were among the main sources of the success stories of England, Holland, and Sweden in the sixteenth and seventeenth century. It is gratifying to be able to say that tolerance pays off."

To conclude, the effects of ethnic diversity appear to be different at different levels of economic analysis (Alesina and La Ferrara 2005). Whereas negative effects on economic outcome are prominent in macro-level studies, the literature reporting a positive effect of ethnic diversity is more micro oriented and focused on regions, cities and firms. In addition, there is a paucity of studies based on primary survey data, in particular of individuals and households. This lacuna is addressed by the present study.

---

[5] INFORM is a EU Horizon 2020 project, Grant no. 693537, conducting multidisciplinary research on formal and informal institutions in the Balkans, and titled: "Closing the gap between formal and informal institutions in the Balkans." Extensive interviews with entrepreneurs throughout this region were implemented in the latter half of 2016. A summary of the main finding is given in Newsletter No. 1:
http://www.formal-informal.eu/files/news/2017/Inform_Newsletter_1_March_2017.pdf

4   The sample, data and identification strategy

In this analysis, we investigate responses from the household sector obtained over the period June-October 2012 from a cross-section survey conducted in Bosnia and Herzegovina by the Prism Research Agency. This agency conducted the "Early Warning Systems surveys" for the United Nations Development Programme (UNDP) in BiH and followed the same procedures according to the same standards to ensure responses representative of the BiH population (UNDPBiH 2000–2010).[6] The interviews were face-to-face meetings based on CAPI methodology, which resulted in no-missing observations.[7]

Although this is a household sector survey, we deal with the individual responses (only one individual per household) of adults who are citizens of BiH. Each individual was randomly selected for interview based on the sample criteria.[8] The targeted sample was 2,000 individuals (we ended up with 2,017 observations) and was designed to be representative of the two entities in BiH (Federation of BiH and Republika Srpska), the District Brcko of BiH, regions (16), municipalities (141), ethnic groups (Bosniaks, Serbs, and Croats), genders, and urban/rural areas. A unique feature of this survey is that it was designed to capture information on ethnic diversity and economic performance at the individual level.

Although the survey dataset has no missing values, "Don't know" or, for some questions, "Don't wish to answer" responses account on average for 3.5% of responses, while around seven per cent of the sample responded in this way to the questions on ethnic minority/majority status and individual earnings. In surveys, "Don't know" or "Don't wish to answer" responses are typically relatively high for income and ethnicity variables. Nonetheless, in this survey, the incidence of these responses to the main questions of interest – on personal incomes and on ethnic self-identification and neighbourhood ethnic diversity – is either below or not much higher than five per cent (see Table 2 below). In the imputation literature, this is the level at which simple "listwise deletion" – i.e. omission – is regarded as unlikely to lead to substantial bias, even when the missing values are not necessarily "missing at random". Nonetheless, we adopt a conservative approach to conserving data by specifying our model with dummy variables for the "Do not know"/"Don't wish to answer" responses of each control variable, while our variable of interest – ethnic diversity (*d_diversity*) – is defined as a binary variable by

---

[6] Prior to the survey implementation, the survey questions were piloted by the researcher in seven cities (40 individuals) in May 2012. The sample covered both entities and the three dominant ethnicities, where each ethnicity was the majority in two cities. After piloting, relevant modifications were made to the questions. The survey questionnaire in English translation is reproduced in Appendix 1.

[7] CAPI refers to Computer-Assisted Personal Interviewing, which is face-to-face interviewing based on computer technology used to administer the questionnaire. According to Foster and McCleery (1999), a major advantage of CAPI is that it reduces respondent error; and routing errors are eliminated because the script automatically routes to the correct questions. In addition, it ensures that data are generally more complete, can considerably reduce the number of "non-responses" and, correspondingly, the need for corrective editing.

[8] To ensure random selection also within the household, the interviewers asked to speak with the adult household member whose birthday was closest to the date they were interviewing.

Table 2: Definitions, construction and descriptive statistics of variables

Variables	Explanation of variables	FULL SAMPLE			HOMOGENOUS AREA $d\_diversity$=0		DIVERSE AREA $d\_diversity$=1		Statistical significance: difference of means between homogenous and diverse sub-samples at the 5% level
	("Do not know/wish to answer" responses reported in the fourth column)	Number of observations	"Do not knows" in %	Mean	Number of observations	Mean	Number of observations	Mean	
earning	Personal income (BAM): 1=0; ...; 11->over 3000	1870	7.3%	3.48	758	3.11	1074	3.75	Yes
fearning	Family income: (BAM): 1=0; ...; 11->over 3000	1755	12.9%	4.79	716	4.48	1004	5.04	Yes
age	Age of respondents in years	2017	0.0%	51.82	825	49.32	1150	53.32	Yes
ed_primary (B)	Primary education: 1-yes; 0-other	2017	0.2%	0.27	825	0.35	1150	0.20	Yes
ed_highschool	Secondary education degree: 1-yes; 0-other	2017	0.2%	0.51	825	0.49	1150	0.54	Yes
ed_collegue	College degree: 1-yes; 0-other	2017	0.2%	0.10	825	0.07	1150	0.11	Yes
ed_university	University degree: 1-yes; 0-other	2017	0.2%	0.11	825	0.09	1150	0.13	Yes
ed_mscphd	Msc or PhD degree: 1-yes; 0-other	2017	0.2%	0.01	825	0.01	1150	0.02	Yes
d_diversity	Ethnic diversity: 1-diverse; 0-homogenous	1975	2.1%	0.58	825	0.00	1150	1.00	-
ethn_minority (B)	Self-identified as ethnic minority: 1-yes; 0-other	2017	7.9%	0.73	825	0.84	1150	0.67	Yes
ethn_majority	Self-identified as ethnic majority: 1-yes; 0-other	2017	7.9%	0.19	825	0.13	1150	0.23	Yes
area_urban	Geographic area: 1-urban; 0-other	2017	0.2%	0.35	825	0.19	1150	0.48	Yes
area_rural (B)	Geographic area: 1-suburban and rural; 0-other	2017	0.2%	0.64	825	0.81	1150	0.52	Yes
d_female	Gender: 1-female; 0-male	2017	0.0%	0.51	825	0.53	1150	0.50	No
marr_single (B)	Marital status: 1-single; 0-other	2017	0.3%	0.17	825	0.19	1150	0.16	Yes
marr_married	Marital status: 1-married; 0-other	2017	0.3%	0.64	825	0.65	1150	0.63	No
marr_cohabitation	Marital status: 1-cohabitation; 0-other	2017	0.3%	0.003	825	0.004	1150	0.003	No
marr_separated	Marital separated: 1-separated; 0-other	2017	0.3%	0.03	825	0.02	1150	0.04	Yes
marr_widow	Marital widow: 1-widow; 0-other	2017	0.3%	0.15	825	0.14	1150	0.15	No
her_dif	Change in the Herfindahl index 2013-1991	2015	0.1%	0.22	823	0.20	1150	0.24	Yes
ethn_friend	Friendship network diversity: 0% to 100%	2017	0.0%	26.77	825	19.10	1150	32.04	Yes
war_migrat_short	War migrants: 1-return shortly afterwards; 0-other	2017	0.2%	0.11	825	0.13	1150	0.10	Yes
war_migrat_years	War migrants: 1-return after years; 0-other	2017	0.2%	0.22	825	0.19	1150	0.24	Yes
war_migrat_noreturn	War migrants: 1-did not return; 0-other	2017	0.2%	0.12	825	0.11	1150	0.14	Yes
war_migrat_no (B)	War migrants: 1-did not migrate at all; 0-other	2017	0.2%	0.52	825	0.55	1150	0.51	Yes
outbih_migr	External migrants: 1-external migrants; 0-other	2017	0.1%	0.07	825	0.06	1150	0.08	Yes
outbih_nomigr (B)	External migrants: 1-non-external migrants; 0-other	2017	0.1%	0.93	825	0.94	1150	0.92	Yes
capitaldistn	Distance of municipalities from Sarajevo in meters/100	2017	0.0%	1.42	825	1.56	1150	1.32	Yes
altituden	Altitude of municipalities/100	2017	0.0%	3.09	825	3.01	1150	3.14	No
abihseven	Locations of headquarters of ABiH corps during the war	2017	0.0%	0.21	825	0.13	1150	0.26	Yes

Regional dummies (16), municipal dummies (127), and "Do not know dummies" (6) omitted for reasons of space.
(B) refers to the base or omitted category in the models reported in Table 3

excluding the "Do not know"/"Don't wish to answer" responses, which simplifies interpretation of marginal effects and our treatment of potential endogeneity. To anticipate, the results reported below are fully robust to the inclusion of a dummy variable for this category of *d_diversity* (Section 7).

The second challenge to valid estimation with survey data is to take account of survey design. The survey design for this sample is straightforward: the whole of BiH was stratified into 16 regions (10 covering the Federation of BiH, five the Republika Srpska and one the District Brcko of BiH); then households (and individuals within households – see Footnote 6) were sampled within each stratum (region). In our case, no Finite Population Correction (FPC) was made. However, the disadvantage of this simple design is that it makes no allowance for possible clustering effects in the data. Consequently, in estimating our model we:

1. ignore the stratification, but instead
2. report cluster-robust standard errors to allow for arbitrary patterns of correlation at the level of the 127 municipalities covered by the sample (from 141, some of which have a population as small as a few hundred), and
3. we include full sets of
   - municipality dummies (with one omitted) and
   - regional dummies, with Sarajevo as the omitted reference category.

This estimation strategy minimises the possibility of omitted variables related to location and adopts a conservative approach to inference.

The ethnic diversity of the surveyed area can be assumed to be exogenous, because in BiH current ethnic composition within the country was primarily war-induced some two decades earlier. This is a most important assumption: if valid, we can identify the effect of ethnic diversity on economic performance and offer a secure platform for policy development. Here, we follow Haller and Eder (2015) who assume that ethnic diversity can be considered as an independent source of the economic performance of individuals and households. Still, in Section 7, we relax the assumption of the exogeneity of ethnic diversity and instrument the variable of interest. First, however, we discuss the validity of our initial exogeneity assumption by addressing two possible sources of endogeneity: namely, omitted influences from the socio-economic environment; and omitted personal characteristics.

There may be characteristics of the socio-economic environment that influence both individual and household economic performance and diversity. If these were to be omitted from the model then their influence would be wrongly attributed to diversity, thereby causing estimates of the diversity effect (if any) to be biased. To control for this possibility, we include dummy variables for urban versus other areas, since urban locations are more attractive to migrants and generally more diverse; hence, economic migration in the post-war period could structurally affect diversity of urban *versus* other areas. Moreover, some regions are more developed than others, host different international organizations and institutions, have a better public infrastructure (e.g. the capital city, Sarajevo), and as such might be more attractive for citizens of all ethnicities, which also could affect diversity. We control for this effect by estimating all models with variables that control for 16 regions in the country, with Sarajevo as the baseline category. Finally, we address potentially unobserved heterogeneities at municipal level by reporting models specified with a full set of municipal dummies (Table 3).

There remains the possibility that the neighbourhood level – to which the survey responses on diversity pertain – has some unobserved characteristics, separate from those at higher levels of aggregation, that both determine economic performance and correlate with diversity. Here, the overriding importance of gaining information on diversity by asking questions in terms that people understood (i.e. pertaining to their neighbourhood rather than to an administrative unit whose boundaries may not be well understood) took priority over our ability to define controls at levels of aggregation matching our survey responses. However, in the context of a population of around 3.5 million, the municipalities have a mean population of just under 25,000 and so aggregate neighbourhoods with at least some strong similarities suitable for control by municipality effects. Given the lack of data currently available at sub-municipality level, this is as far as we can take this discussion. However, to anticipate our results reported below, we found that inclusion of municipality dummies makes no noteworthy difference to the estimated effects of diversity or of any of the control variables (Table 3). This suggests that omission of dummies at still lower levels of aggregation is unlikely to be a source of substantial bias.

In addition, for each of the 127 municipalities we construct a measure of the change in the ethnic diversity concentration by calculating the Herfindahl index for all ethnic groups (1991 and 2013). We then calculate the difference between the Herfindahl index from the last pre-conflict census in BiH (1991) and the data from the first post-conflict census (2013) to obtain the change in ethnic diversity proxied by this measure. Accordingly, zero indicates no change; negative values indicate increase in the ethnic diversity; and positive values indicate increase in the ethnic homogeneity. The average value of this index is 0.22, ranging between -0.19 to 0.47, suggesting an average increase in homogeneity in the sample. The correspondingly right-skewed distribution of changes in the Herfindahl index is reported in Figure 1.

We assume that the whole range of outcomes reflects conflict-induced forced migration: not only were decisions to move determined by ethnicity rather than by the range of economic and other influences typically influencing migration; but also decisions to stay, where belonging to the largest ethnic group offered "safety". By including this index in our models we control for the variability between those (few) municipalities experiencing little change between pre- and post-conflict times and the many experiencing substantial shifts away from ethnic diversity and towards ethnic homogeneity. In doing so, we also address a potential source of omitted variables bias in the case that previous changes in ethnic diversity are not only related to the current levels of perceived ethnic diversity but also to current levels of economic performance.

Potentially omitted personal characteristics are a corollary of both limits to the feasible length of our questionnaire (with respect to both cost and the patience of respondents) and an inherent limitation of cross-section data (precluding the use of individual fixed effects to control for unobserved and unobservable time invariant influences). Accordingly, to gain an indication of the possibility that our estimates of the income effects of ethnic diversity are unduly influenced by omitted variables bias, we investigate estimates from successively more developed model specifications. To anticipate, we find that the estimated diversity effects are stable with respect to increasing numbers of control variables and supplementary diversity variables. While not definitive, this approach at least provides some reassurance that our variable of interest is not simply capturing the influence of personal characteristics omitted from our models.

Figure 1: Changes in the Herfindahl index of ethnic diversity in BiH, 1991-2013 (red line presents the mean value) (percentages of municipalities)
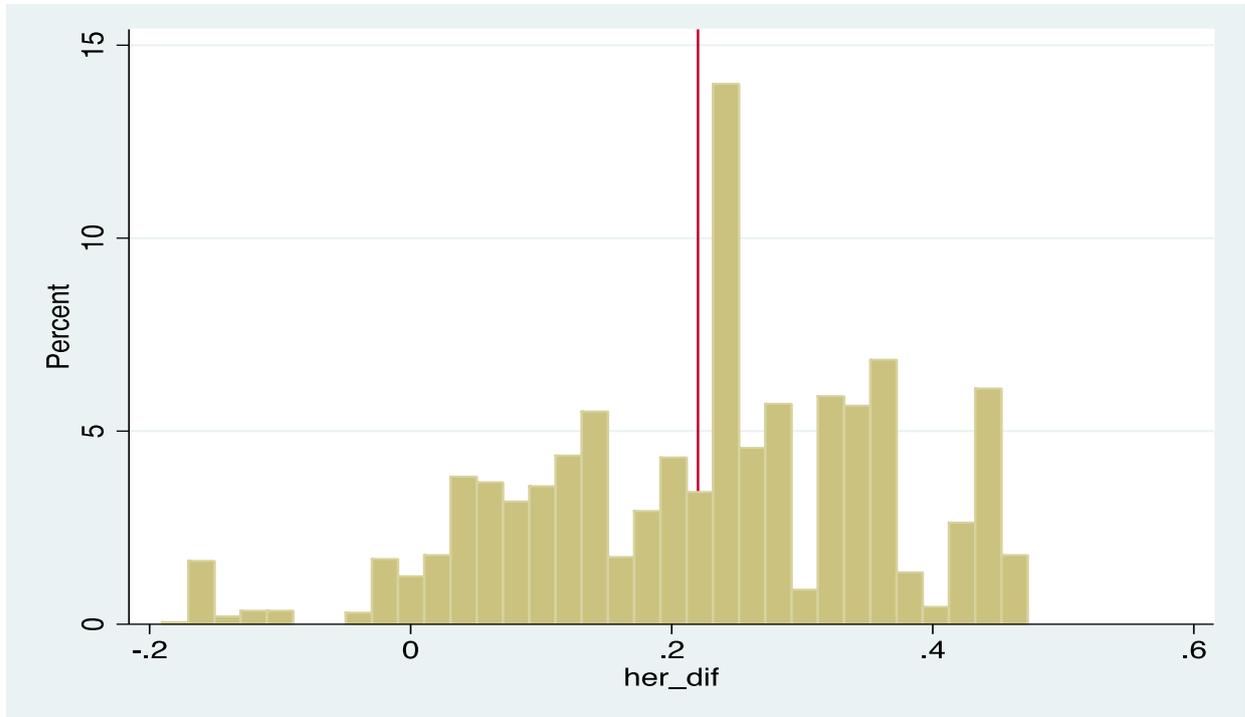
Source: authors.

Based on most findings by the diversity literature at the micro level, we expect ethnic diversity to have beneficial effects on individual economic performance. Accordingly, we state our main hypothesis:

H0: Ethnic diversity in BiH is beneficial for the incomes of individuals and their families.

5   Variables and model specification

The specification of our model is guided by theoretical reasoning concerning the potential links between ethnic diversity and the economic performance of individuals and households (families) measured by their levels of total monthly income. Our theoretical starting point is the Mincer earnings function (Mincer 1974) that relates earnings to schooling and experience. This approach to modelling earnings remains one of the most widely used specifications in empirical economics (Lemieux 2006), and supports a wide range of augmented models. Accordingly, in each of the models estimated below, respondents' personal income – *earning* – is a function of (*f*) respondents' age (*age*), squared age (age^2) and five categories of education – primary education, secondary education, college degree, university degree, and MSc/PhD degree (*ed_primary, ed_highschool, ed_college*, *ed_university, ed_mscphd*). These "Mincer variables"

are augmented by: (i) our variable of interest, a dummy variable *d_diversity*, capturing the effect of whether the respondents' neighbourhoods are ethnically homogenous (0) or diverse (1) on respondents' personal (and, later, family) income; and (ii) a variety of control variables.

Personal/family income = f (Mincer variables; *ethnic diversity*; *control variables*)   (1)

Table 2 provides definitions, construction details and descriptive statistics for the dependent and all independent variables in the full sample as well as in two subsamples differentiated by the variable of interest (*d_diversity*): i.e. ethnically homogenous neighbourhoods (*d_diversity* =0); and ethnically diverse neighbourhoods (*d_diversity* =1). The final column reports, for each variable, the outcome of a t-test of the null that there is no difference between the sub-sample means. Accordingly, the unconditional means show that in the more diverse neighbourhoods, compared to the more homogeneous, respondents have higher personal and family earnings, are older, better educated, more urban, have more experience of ethnic change, and are more likely to have friends from other ethnic groups.

The dependent variable in our modelling is the economic performance of individuals and families, which we proxy by the total monthly income of respondents. BiH is a country with huge unemployment (based on ILO definition - International Labour Organization – it would be around 28% in the year of our survey, 2012) and, correspondingly, a substantial amount of income is earned in the grey economy, which is estimated to be around 30% of total economic activity. (Our survey indicates a similar percentage of informally employed respondents.) Accordingly, in order to capture better the economic performance of individuals, we asked participants to report their total monthly income from whatever source, and provided them with a scale of different income categories. While some previous studies in the ethnic diversity literature use official indicators to proxy the productivity of individuals (e.g. Ottivano and Peri 2006; Bellini et al. 2012) our approach, which includes earnings from the informal economy, is more appropriate for BiH.

The income scale used in this survey (Appendix 1) had previously been tested and validated by the professional research agency conducting the survey for this project (UNDPBiH 2000–2010). Reflecting the particular circumstances of BiH, the scale has 11 categories, ranging from "No income in the previous month" to "over 3,000 BAM in the previous month" as the largest category. The scale differs from typical practice in more developed economies by starting with "No income", an approach validated by 23% of respondents reporting no personal income and 9% reporting no family income. The responses are right skewed, with more than 70% of responses recorded for the first four categories of personal income and over 50% in the case of family income (Appendix 2). This distribution affects the choice of our estimator. First, we reject the common practice of taking category mid-points and treating our categorical variable as if it were continuous. Not only would this approach entail arbitrary top coding (a practice that is methodologically dubious in general) but this approach would also lose the important information that the poorest individuals and households have no regular income but are very likely to be dependent on the goodwill of family and friends (which would introduce a particular distortion into the present study). Second, we cannot use interval regression, because this model assumes normality, whereas the distribution of incomes is nonnormal. The usual procedure would thus be to model the log of wages by transforming the upper and lower bounds of each

category into natural logarithms. However, this procedure would entail the loss of the zero observations, which – as we have noted – are particularly important in this investigation. The alternative is to model our data by estimating an ordered probit (OP) model.[9] OP generalizes the probit model to analyse a dependent variable defined as 'a finite set of ranges, each corresponding to one possible outcome' and allowing for cut points defining uneven ranges (Roodman 2011: 164). Accordingly, our estimation strategy proceeds as follows: first, we estimate an OP model in which our variable of interest (ethnic diversity, *d_diversity*) is treated as exogenous; but, second, we allow for the potential endogeneity of *d_diversity* by estimating an instrumental variables (IV) version of OP available in the Stata user-written program (*cmp*) (Roodman 2011).

Measuring ethnic diversity is not a straightforward task and it seems that researchers are still searching for an appropriate measure or indicator of ethnicity and diversity. Generally, ethnicity and ethnic groups are slippery concepts (Fearon 2003) and far from straightforward (Posner 2000). Posner (2000), Constant et al. (2009) and Efendic at al. (2011; 2015) use different measures of ethnic diversity obtained from survey data. The survey approach seems promising for our research, bearing in mind that ethnic perceptions are subjective and the product of self-definition (Posner 2000).[10]

There are other, country-specific reasons why responses to our survey on ethnic self-identification and perception of ethnic diversity might be more valid than measures from the official census. In the census data, the ethnic composition of municipalities is biased towards reflecting the ethnic composition of people with a right to register in a particular municipality rather than those actually living there, which is the measure of interest in the present study. Because the BiH diaspora (the largest in Europe, in relative terms) tended to register, together with some internal migrants, in their municipality of origin, the census data published in 2016 does not necessarily reflect the reality of the ethnic composition on the ground. Using census data to measure ethnic diversity is further confused by publication in slightly different versions; although the census data are published by the national statistical agency, the statistical authorities from one entity (RSBiH) published their own entity data, which yields a slightly different picture with respect to ethnic structure. Taking into account these issues with the census data on ethnic diversity, we believe that our perception measure of ethnic diversity – derived from a nationally representative survey – gives a more accurate picture of the ethnic diversity actually experienced by people.

To get a measure of diversity, we asked respondents to describe the level of ethnic diversity in their neighbourhood. Following Efendic et al. (2015), we provided a range of possibilities (1–5): from one ethnicity only in the neighbourhood (1); to the area is very diverse (4); while 5 codes "Do not know" responses. We rely on a dummy variable (*d_diversity*) since the responses are right skewed; 41% of respondents claim to live in a fully homogenous area (category 1),

_____

[9] For a particularly clear example of this decision procedure, see StataCorp (2015), 1056–1065 (the entry on *intreg* – interval regression).

[10] Bellini et al. (2012: 8) cite Fearon (2003) to argue that, in ethnology, the "right list" of groups would be based on a process of "self-categorisation" where people recognize the distinction of groups and anticipate that significant actions are or could be conditioned on belonging or not to a group. Accordingly, a direct approach to ethnic identification would involve carrying out worldwide surveys.

while some 31% live in areas with "small" diversity (category 2). Accordingly, *d_diversity* is constructed as a dummy variable: =1 (diversity) for categories 2, 3 and 4; =0 (homogeneity) for category 1.[11] After presenting the main variables of interest, we continue with the model specification and introduce the other control variables.

We include control variables that might be important influences on earnings, including (with omitted or base categories indicated in Table 2): *area_urban*, a dummy variable capturing whether the surveyed area is urban or not (official indicators for BiH suggest that average earnings are generally higher in urban areas); *d_female* coding the gender of respondents (official sources also report that gross earnings of male respondents are higher than of female respondents); and *marr_married, marr_separated, marr_cohabitation,* and *marr_widow*, respectively coding whether respondents are married, separated, widowed, or live in cohabitation, while the omitted category is if the respondent is single (important in the distinction between our individual and household models, and because 'married men are expected to work more years over their lifetime than married women and hence have higher wages while for single men and women wages are roughly similar'; Ibrahim 2017: 56); and *ethn_majority*, a dummy indicating self-identification of respondents as belonging to the majority ethnic group in comparison to a minority ethnic group (in the area where they live). Since BiH is a post-ethnic conflict country, we capture ethnic status according to whether individuals perceive themselves as belonging to a minority or to the majority ethnic group. Fearon (2003) reports that around 70 per cent of countries in the world have an ethnic group that accounts for the absolute majority; however, that is not the case for BiH. In the context of BiH it is important to control for majority/minority status within the country, not least because each of the three constituent ethnicities appear to be in some parts of the country a majority while, in others, a minority. Accordingly, at the individual level, there may be variations in economic variables depending on self-reported majority/minority status, which is often identified as important in empirical research (e.g. Efendic 2010; Vanhoutte and Hooghe 2012). If there is discrimination based on ethnic minority-majority status (e.g. minorities sometimes experience more problems in finding employment, hence might have systematically lower income), this variable should capture this effect. Finally, we include dummy variables for each surveyed region (there are 16 regions that we control for in all specifications) and municipality (there are 127 municipalities), as well as "Do not know" dummies (six in the fully specified model) for control variables having such responses. We have explained the purpose of these variables in the previous section.

---

[11] Estimation of our preferred model with separate dummy variables for each category, including "Do not know", and omitting Category 1 as the base, yields a statistically significant estimate only for Category 2 (p=0.003). Post estimation testing reveals that the Category 3 and 4 effects are not strongly statistically different from one another (p=0.080), while the Category 4 estimate is not significantly different from the Category 2 estimate (p=0.364). In addition, Category 5, "Do not know", yields a statistically insignificant result (p=0.878). Accordingly, we combine Categories 2, 3 and 4 and omit Category 5 to define our variable of interest in binary terms.

6   Estimation and discussion of results

We estimate our initial single-equation model by the Ordered Probit method, with cluster-robust standard errors, clustering on municipalities. The results of this specification and all other augmented models used to investigate the stability of the ethnic diversity effect are reported in Table 3. Greene (2012: 828–830) explains that in OP 'the partial effects of the regressors (**x**) on the probabilities are not equal to the coefficients'. Hence, 'we must be very careful in interpreting the coefficients … Indeed, without a fair amount of extra calculation, it is quite unclear how the coefficients in the OP model should be interpreted.' Our main purpose is to estimate the marginal effect of living in an ethnically heterogeneous neighbourhood rather than in an ethnically homogenous neighbourhood on the probabilities that our respondents will be in a particular income category. Table 4 reports these estimated marginal effects, which are derived from the estimates reported in Table 3. For reasons of space, we do not report the derived marginal effects for the other variables.

In all models for personal income (Table 3, Models 1–7) each of the ten cut points is estimated with a high level of precision ($p<0.05$); in our family income model (Model 8) two cut points are not precisely estimated. This validates the intervals of our income categories. (The cut points are not reported for reasons of space.)

Diagnostic testing for nonlinear models is still not much developed (UCLA 2015), and this includes the lack of any established procedure for choosing between models. One diagnostic procedure available to researchers is testing the "parallel regression assumption" (Wooldridge 2010: 658; Greene 2012: 831–832), which describes the assumed equality of the slope coefficients across all the ordered categories of the dependent variable (i.e. across all outcomes). Although Williams (2006) suggests that common practice is to ignore potential violations, we test this assumption by comparing two models:[12] (i) a restricted model in which the estimated coefficient on our variable of interest – *d_diversity* – is constrained to be constant across all income categories while all other coefficients are free to vary across income categories; and (ii) a fully unrestricted model in which all coefficients are free to vary (including *d_diversity*). A likelihood ratio test was unable to reject the null hypothesis that the unrestricted model adds no explanatory power to the restricted model, which is consistent with the parallel regression assumption ($p=0.61$). Bearing in mind that both models make extreme demands on the data – as every unrestricted coefficient must be estimated for each category – we could implement this test only for Model 2 reported in Table 3 and then with education captured by a single categorical variable rather than separate dummies for each category. (None of the more fully specified models could converge to a solution.) Nonetheless, given that the estimated slope coefficient on *d_diversity* is statistically significant and stable across all the Ordered probit models reported in Table 3, we conclude that it is reasonable to assume parallel regressions for our variable of interest. Nonetheless, because we were unable to test all the variables for

_____

[12] We use the Stata user-written programme *gologit* (Williams 2006) to estimate the more and less restricted models as generalized ordered logit models and then compare them by means of the likelihood ratio test. Our procedure is based on a suggestion by Williams in a public discussion: https://www.statalist.org/forums/forum/general-stata-discussion/general/783-diagnostic-tests-in-a-ordered-probit

Table 3: Ordered Probit (OP), heteroskedastic OP, and IV OP models

Variables	ORDERED PROBIT MODELS (z-statistics calculated from cluster-robust standard errors; clustered on municipalities; henceforth: cluster-robust)					HETEROSKEDASTIC ORDERED PROBIT cluster-robust	IV ORDERED PROBIT MODELS cluster-robust	OP FAMILY INCOME MODEL cluster-robust
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.	Coeff.
age	0.05***	0.05***	0.04***	0.04***	0.04***	0.02**	0.04***	-0.02*
age2	-0.0003***	-0.0004***	-0.0004***	-0.0003***	-0.0003***	-0.0001**	-0.0003***	0.00006
ed_highschool	0.60***	0.56***	0.48***	0.47***	0.46***	0.14**	0.42***	0.42***
ed_collegue	1.15***	1.10***	1.00***	1.01***	0.99***	0.33***	0.91***	0.93***
ed_university	1.40***	1.35***	1.25***	1.26***	1.24***	0.48***	1.16***	1.23***
ed_mscphd	1.91***	1.86***	1.72***	1.74***	1.78***	0.72***	1.69***	1.56***
d_diversity	-	0.18***	0.17***	0.17***	0.17***	0.08**	0.51**	0.20***
area_urban	-	-	0.13**	0.14**	0.14**	0.05*	0.05	0.31***
d_female	-	-	-0.35***	-0.38***	-0.39***	-0.14***	-0.35***	-0.04
ethn_majority	-	-	0.12*	0.08	0.08	0.03	0.16**	0.17**
marr_single	-	-	-0.13*	-0.18*	-0.17	0.02	-0.12	-0.11
marr_cohabitation			0.90**	1.00***	1.01***	0.36*	0.98**	0.73
marr_separated			0.02	0.06	0.08	0.05	-0.003	-0.29**
marr_widow			0.26***	0.27***	0.25***	0.09**	0.24***	-0.31***
her_dif					0.32**	0.03	-0.03	-0.009
ethnic_friend	-	-	-		-0.0001	-0.000006	-0.0007	0.0001
war_migrat_short	-	-	-		-0.01	-0.02	-0.05	0.10
war_migrat_years	-	-	-		0.11	0.04	0.01	-0.06
war_migrat_noreturn	-	-	-		0.19*	0.08*	0.13	0.15

outbih_migr	-	-	-		0.07	0.03	0.03	0.11
Do not knows dummies (1-6)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Regional effects (16)	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Municipal effects (127)	Yes	Yes	No	Yes	Yes	Yes	No	Yes
IV Instruments (3)	No	No	No	No	No	No	Yes	No
Number of observations	1,870	1,832	1,832	1,832	1,831	1,831	1,831	1,719
Statistical significance: 1% denoted as ***, 5% as **, 10% as *.								
Model 1: Mincer's base specification OP; **Model 2:** Mincer's specification with ethnic diversity effect OP; **Model 3:** Diversity personal income model without municipal dummies OP; **Model 4:** Diversity personal income model with municipal dummies OP; **Model 5:** Fully specified diversity personal income model OP; **Model 6:** Fully specified diversity personal income model – heteroskedasticity robust OP; **Model 7:** Fully specified diversity personal income model – IV OP; **Model 8:** Fully specified diversity family income model OP.								

coefficient equality, we do not interpret and discuss either the "Mincer variables" – age, age-squared and education – or the other control variables, although it is apparent that their effects are also stable across the different specifications reported in Table 3 and that they are mainly estimated with acceptable levels of statistical significance. Instead, we focus on our variable of interest.

Model 4 differs from Model 3 only by inclusion of 126 municipal dummies (the remaining one is omitted as the base category). It is striking that for practical purposes the results are identical. Our main finding is that respondents living in more diverse areas, on average and holding all other factors constant, have higher probability of being in a higher income category in comparison to those individuals living in ethnically homogenous areas. Moreover, this finding holds for family income as well (Model 5).

7   Robustness of the reported results

We check the robustness of our results in three ways: (i) by augmenting our ordered probit models with additional controls; (ii) by estimating alternative models that both correct for heteroskedasticity and allow for a distributional assumption different from the one imposed by ordered probit; and (iii) by estimating an instrumental variables approach to ordered probit models to address the potential endogeneity of our variable of interest.

First, we augment Model 4 with two new dimensions of ethnic diversity. Models 5–7 include our Hefindahl index of change in ethnic diversity 1991-2013 for each municipality (*her_dif*, described above). The purpose of this variable is to control for the possibility that economic outcomes are influenced not only by current ethnic diversity but also by changes in the recent past. Together with this variable, we include a variable to indicate whether respondents' friendship networks are diverse or homogeneous (*ethnic_friend*). This is intended to control for the possibility that what matters for economic outcomes may be individual preferences expressed in the choice of friends rather than the ethnic diversity in respondents' social environment.

Second, although we argue that current patterns of ethnic diversity in BiH are predominantly the result of conflict-induced migration, post-conflict return migration may have had an additional influence on the patterns of ethnic diversity observed by our survey. Migrants who spent some time abroad and returned might be a source of new capital, ideas, know-how, and social capital in the form of diaspora networking, or even a source of additional income in the form of remittances (Efendic 2016). In turn, these may influence personal and family incomes. To investigate this possibility, we augment our model with four variables to capture different categories and, hence, potentially different motivations, of migrants: compared to the approximately 50 per cent who did not leave their homes during the war, we estimate the effects on income of being forced out of home but returning from internal or external migration (i) almost immediately (*war_migrat_short*), (ii) after a few years (*war_migrat_years*), and (iii) never returning to their home (*war_migrat_noreturn*); and (iv) post-conflict migrants who spent up to a year or more abroad (i.e. not likely to be tourists) in the last three years prior to the

survey (*out_bih_migrat*).[13] The base categories are indicated in Table 2; "Do not know" responses are also included.

Model 5 reported in Table 3 demonstrates that inclusion of these additional controls makes little difference to our estimated diversity effects. Nevertheless, these findings are consistent with our argument that the first-order determinant of current patterns of ethnic diversity was violence and the threat of violence, suggesting that our variable of interest may be regarded as exogenous. Our investigation of marginal effects will be based on these "fully specified models".

Heteroskedasticity, when the variance of the regression residual varies systematically across observations, may invalidate both statistical inference and point estimates from ordered models. Accordingly, our second robustness check was to estimate "heteroskedastic ordered models" (also known as "heterogeneous choice models") that 'explicitly specify the determinants of heteroskedasticity in an attempt to correct for it' (Williams 2010: 540).[14] These models simultaneously fit two equations, one for the outcome and one for the residual variance, thereby allowing the variance to differ across observations, 'hence correcting for heteroskedasticity' (Williams 2010: 543). In our application, Model 6 replicates the full specification of Model 5, while allowing the variance to be influenced by all of the independent variables except the municipal and "Do not know" dummies (convergence could not be achieved when these were included). For reasons of space, we do not report the variance models. The "usual suspects" proved to be statistically significant determinants of the residual variance – e.g. being female (reduced variance), having university education (increased variance), being an external migrant (increased variance) and being single or widowed (reduced variance) – but these did not include our variable of interest (*d_diversity*). In the outcome equations reported in Table 3, the estimated coefficients on our variable of interest continue to be positive and statistically significant, although they differ in size. In Model 6 the estimated effect of *d_diversity* is smaller than in the corresponding OP model (0.08), although the derived marginal effects are similar (see below). We conclude that at least our qualitative conclusions so far are not an artefact of misspecification. The quantitative implications of these estimated coefficients are discussed below when we analyze the marginal effects.

We also use a heteroskedastic ordered model to investigate the validity of the transformation applied to our dependent variable by specifying our model with an alternative link function. Williams (2010: 546), quoting Norušis (2005: 84), suggests that probit models 'are reasonable choices when the changes in the cumulative probabilities are gradual', which applies reasonably well to our dependent variable. Yet Stata's complementary log-log model (*cloglog*) can be appropriate when 'the cumulative probability for lower scores is high and the approach to 1 is slow', which also applies reasonably well to our dependent variable. Our fully-specified personal income model (Table 3, Model 5) would not converge using the cloglog link.

_____

[13] This question asked respondents if they spent this time outside of the countries of Former Yugoslavia to avoid confusing external migration with the common practice of Bosnian Croats and Bosnian Serbs possessing dual citizenships with Croatia and Serbia respectively spending periods in these countries as residents. This question thus captures external migrants who spent some time abroad, mainly in Western Europe, which is the most attractive region for external migrants from the Western Balkans.

[14] Implemented in Stata 14 by the user-written programme *oglm* (Williams 2010).

Accordingly, we omitted the municipal dummies and compared the smaller model estimated with *probit* and *cloglog* links respectively. However, the two sets of derived marginal effects revealed no noteworthy differences in either magnitude (to the third decimal place) or levels of statistical significance. Accordingly, this procedure revealed no evidence conflicting with our choice of ordered probit.

As the next check, we estimate our preferred model (Model 5) controlling for "Do not know" responses of the variable of interest. The obtained results for the variable of interest remain fully consistent in terms of the estimated magnitude (0.17) and statistical significance (p-value=0.010), while the "Do not know" dummy is statistically insignificant (p-value=0.906).

In our final robustness check, we acknowledge that the assumption of exogeneity of our ethnic diversity variable can be challenged. Accordingly, we use an instrumental variables approach to estimate the effect of our ethnic diversity variable.[15] We are concerned about two sources of endogeneity: (i) simultaneity, or reverse causality, whereby income levels might have affected patterns of migration and, in particular, return migration, thereby influencing ethnic diversity; and (ii) the possibility of unobserved influences on income that are also correlated with ethnic diversity. (We have tried to control for such omitted variables by including many control variables, including regional and municipality dummies, but the limitations of cross-section data preclude certainty in this regard.) Potential instruments must be highly correlated with our potentially endogenous variable, ethnic diversity, while being neither a direct influence on income nor correlated with systematic influences that are unobserved and thus in the error term. Such instruments had to come from outside our dataset. In our search for instruments, we reasoned that "Altitude of the municipality" (*altituden*), "Distance of the municipality from the capital city" (*capitaldistn*) and the "Location of army bases during the conflict" (*abihseven*) could all have affected the intensity of inter-ethnic conflict and thus the pattern of ethnic diversity but have no strong, theoretically motivated relationships with the distribution of individual incomes in BiH. (For example, while, elsewhere, regions of high altitude might be regions of low income, in BiH the Sarajevo region is among both the richest and those with the highest altitudes.)[16] The IV version of ordered probit was estimated by limited-information

_____

[15] We thank the supervising co-editor of this journal who inspired us to identify relevant instruments for this checking procedure.

[16] In brief, these instruments are motivated as follows. (i) *altituden.* Geographical features could influence the nature of the conflict. We investigated a number of candidate geographical instruments and found that the altitudes of municipalities are inversely related to ethnic diversity. We conjecture that this effect worked through its influence on military tactics, especially in the beginning of the war, whereby attacks were typically made from high ground on the more ethnically diverse urban communities typically occupying lower ground. Even the siege of Sarajevo, which – as noted, is a high altitude region – was conducted from surrounding high ground. (ii) *capitaldistn.* As Sarajevo was the centre of a multi-ethnic civil government of multicultural BiH, the distance from the capital city during the war might also have influenced diversity in different locations, which in many cases did not even have a physical connection with Sarajevo. (iii) *abihseven.* Very soon after BiH became an independent state recognized by the UN in the beginning of 1992, the war started, and in April 1992 BiH established its Army, the Army of the Republic of Bosnia and Herzegovina (ABiH). The ABiH was organized in seven corps located in seven municipalities (Sarajevo, Mostar, Tuzla, Zenica, Travnik, Konjic, and Bihac). This was a multi-ethnic army, the one officially fighting for an independent, multi-cultural, multi-ethnic BiH, and the one being ethnically diverse to some extent, including even the highest-ranking officers. This might help to explain why ethnic diversity is better preserved in some regions today than in others, depending on the location of these seven army corps. In line with this reasoning, we find that location of ABiH corps had a positive effect on ethnic diversity while distance from the capital city and altitude had a negative

maximum likelihood (LIML) using *cmp*, a user written programme implanted in Stata 14 (Roodman 2011). Table 3 reports the IV OP results in Model 7).

Model 7 suggests that the effect of ethnic diversity is still positive, statistically significant at (about) the conventional five percent level (p-value=0.051), but considerably larger. This applies to the estimates in Table 3 but also (to anticipate) to the marginal effects reported in Table 4 and Figure 2. Nonetheless, we have two sets of reasons for preferring our more modest – but still economically substantial – OP estimates: first, some issues with IV estimation in general as well as with our particular applications; and, most important, the arguments set out above for regarding ethnic diversity in BiH as exogenous.

First, our instruments are all – of necessity – measured at the level of the municipality whereas we measure ethnic diversity at the neighborhood level, a lower level of aggregation. Moreover, if we specify our models with a full set of municipality dummies then IV OP cannot estimate the model, because convergence cannot be achieved. Although the municipality dummies make little difference to our OP estimates of the diversity effect (compare Models 3 and 4), it is uncomfortable not to be able to include them in our IV specification. Given that IV estimation is 'never unbiased' (Wooldridge 2010: 107), these additional issues mean that the larger diversity effects reported in Table 3, Table 4 and Figure 2 (below) may be subject to considerable bias. Second, and most important, is the overriding consideration that patterns of ethnic diversity in BiH reflect conflict-induced migration and hence – as we argue above – may be regarded as exogenous for analytic purposes. We recognize the possibility of unobserved but systematic influences on post-conflict return migration that might invalidate our exogeneity assumption. However, as we note above, we address this possibility by augmenting our model with three variables to capture different categories and, hence, potentially different motivations, of migrants forced out of their homes by war together with a fourth to capture the effects of returnees from post-conflict migration. Inclusion of these additional controls in our models estimated by OP makes no noteworthy difference to our estimated diversity effects, which is consistent with our argument that current patterns of ethnic diversity may be treated as exogenous. Accordingly, we prefer to highlight the OP estimates reported in Table 4. Even without the caveat that these may well be downwardly biased, our evidence that living in an ethnically diverse neighborhood has economic advantages is sufficiently substantial to be relevant for policy.

8   Quantitative effects

The interpretable quantitative results of this investigation are reported in Table 4 (below) and in Figure 2 (below). Our preferred results are from the fully-specified OP models in which our variable of interest – *d_diversity* – is treated as exogenous (Columns 1–2). The marginal

---

effect on diversity (Table 3). Each of these instruments proved to be statistically significant in the first stage regression.

Table 4: Estimated marginal effects of ethnic diversity (*d_diversity*) derived from fully specified Ordered Probit (OP), heteroskedastic OP, and IV OP models

Variables	PERSONAL INCOME						FAMILY INCOME	
	Fully specified OP model (Model 5)		Fully specified heteroskedastic OP model (Model 6)		Fully specified IV OP model (Model 7)		Fully specified OP model (Model 8)	
	1.		2.		3.		4.	
	Marginal effects	p-values	Marginal effects	p-values	Marginal effects	p-values	Marginal effects	p-values
Income category 1	-0.045	0.007	-0.059	0.003	-0.142	0.047	-0.027	0.003
Income category 2	-0.003	0.014	-0.003	0.046	-0.008	0.019	-0.003	0.007
Income category 3	-0.008	0.012	0.003	0.762	-0.023	0.049	-0.029	0.004
Income category 4	0.010	0.011	0.021	0.007	0.029	0.019	-0.005	0.006
Income category 5	0.012	0.009	0.017	0.003	0.037	0.027	0.006	0.003
Income category 6	0.010	0.008	0.010	0.011	0.030	0.042	0.009	0.004
Income category 7	0.009	0.009	0.007	0.044	0.028	0.057	0.012	0.005
Income category 8	0.007	0.009	0.004	0.205	0.022	0.069	0.013	0.004
Income category 9	0.004	0.016	0.001	0.474	0.015	0.086	0.012	0.003
Income category 10	0.001	0.041	0.0003	0.721	0.005	0.172	0.007	0.006
Income category 11	0.002	0.042	-0.0001	0.971	0.007	0.152	0.006	0.010
Do not knows included (6)	Yes		Yes		Yes		Yes	
Regional effects included (16):	Yes		Yes		Yes		Yes	
Municipal effects included (127):	Yes		Yes		No		Yes	
Instruments included (3)	No		No		Yes		No	
Number of observations	1,831		1,831		1,831		1,719	
Effect of ethnic diversity*	25%		22%		47%		20%	

*Effect of ethnic diversity on mean income if BiH were to change from a counter-factual completely homogenous to a counter-factual completely heterogenous environment. For the personal income calculation, we use the first seven of the 11 income categories, capturing 95% of observations; the family income calculation uses all 11 categories, as income is distributed more evenly across the higher categories.

Figure 2: Estimated marginal effects derived from fully specified Ordered Probit (OP), heteroskedastic OP, and IV OP models – visual interpretation
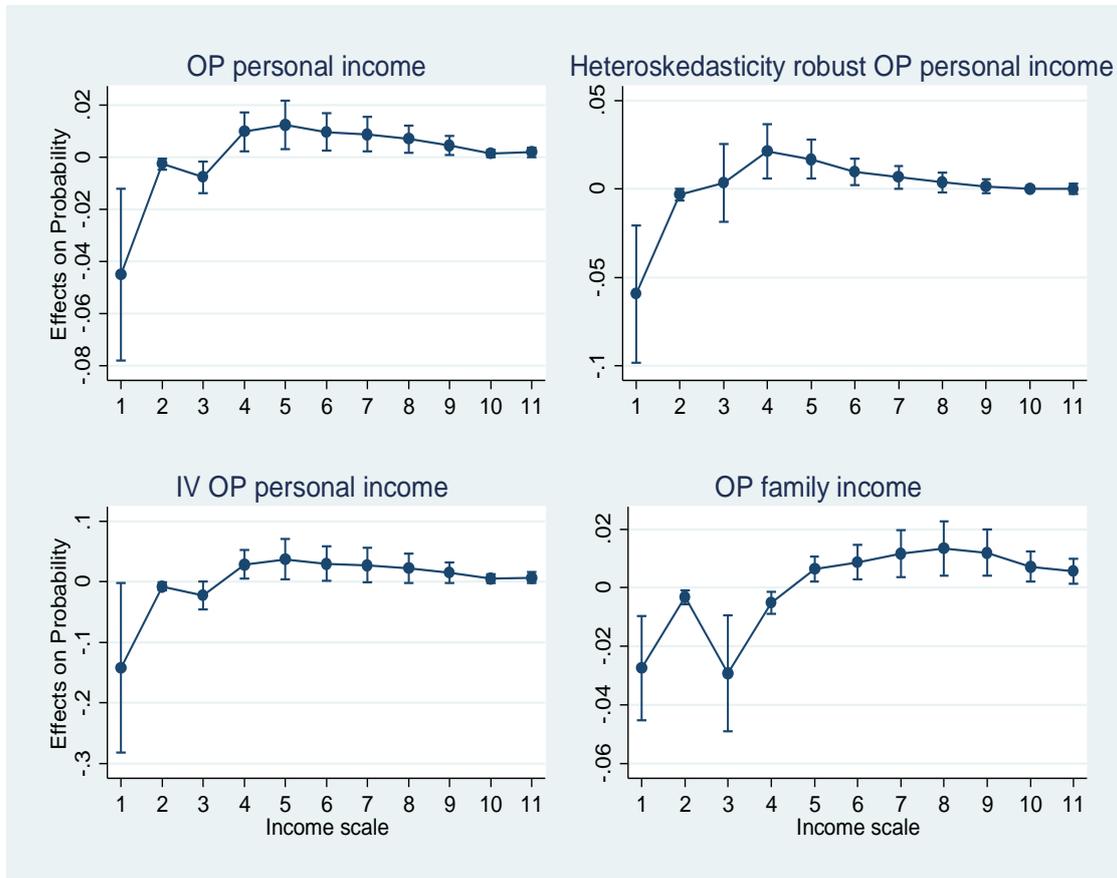
Note. Dots indicate point estimates and vertical bars the corresponding 95% confidence intervals (CIs). A CI overlapping zero indicates that the corresponding point estimate is not statistically significant at the conventional five per cent level.

effect for Income Category 1, holding all other influences constant, means that living in an ethnically diverse area rather an ethnically homogeneous area reduced respondents' chances of having no income in the previous month by 4.5 percentage points. For zero family income in the previous month (Column 4), the corresponding reduction is 2.7 percentage points. For comparison, the heteroskedastic ordered models suggest larger effects – reductions of six percentage points (Column 2); while the IV model suggests still larger effects – 14 percentage points (Column 3). Income Categories 2 and 3 generally record negative but smaller effects. Together, these negative effects suggest that the effect of ethnic diversity in respondents' neighbourhoods is to reduce their chances of being in the lowest income categories. This has implications for poverty reduction strategy. In contrast, ethnic diversity increases the probability of being in Income Category 4 (in all personal income models), and of being in all higher Income Categories across all models. For example, for Income Category 5, the OP

models reported in Columns 1 and 2 suggest 1.2 and 1.7 percentage points improvement in the probability for personal incomes and a 0.6 percentage point improvement for family incomes. Again, the IV model suggests larger effects for personal income (3.7 percentage points in Column 3). Overall, this evidence suggest that the effect of local ethnic diversity is to reduce the probability of low income (Categories 1, 2 and 3) and increase the probability of medium and high incomes (Categories 4 and above, with the caveat that the very highest categories – 10 and 11 – account for very few respondents and are not estimated in all models at conventional levels of statistical significance). The overall conclusion is illustrated in Figure 2, which plots the marginal effects reported in Table 4. Qualitatively, the panels reveal the same pattern for personal incomes from all three estimation strategies: reductions in the probability of zero income in the previous month and very low income levels (i.e. poverty reduction); and increases in the probabilities of being in higher income categories. The bottom right panel reveals much the same pattern for family incomes.

Finally, we indicate the overall quantitative effect of ethnic diversity on mean income. The marginal effects reported in Table 4 are calculated by taking the difference between two counter-factual states: namely, the probability of respondents being in each income category if every neighborhood were to be ethnically homogeneous (*d_diversity*=0); and the probability of respondents being in each income category if every neighborhood were to be ethnically diverse (*d_diversity*=1). We used these estimated probabilities to weight the mean income in each of the first seven income categories (omitting the four highest, which include few observations and are not estimated with statistical significance in all models). We then summed these weighted means to obtain the implied overall mean income under both ethnic homogeneity and ethnic diversity. The percentage differences for each model are reported in Table 4. (These calculations for the preferred personal income models are presented in full in Appendix 2.) From our preferred OP models, treating local diversity as exogenous, we find that the average effect of ethnic diversity on mean income in our OP model is an increase of around 20% for both personal and family income. Although these are our lowest estimates, the magnitudes are sufficiently large to be of both scientific and policy interest.

To conclude, the empirical evidence that we report in this study does not reject our maintained hypothesis. Our estimates suggest that the effect of ethnic diversity has an economically positive effect, which should not be ignored by policy makers. Finally, the validity of this finding is suggested by its consistency with Efendic et al. (2015) who report a positive effect of ethnic diversity in the business sector of the BiH economy.

9   Conclusion

We collected nationally representative cross-sectional survey data to investigate the effect of ethnic diversity on individual and household incomes in Bosnia and Herzegovina, a post-conflict country. The literature review establishes that ethnic diversity can in theory have positive or negative effects on economic performance, and that different findings in this respect might be associated with the sample and/or level of analysis (micro or macro). At the micro level, the literature suggests that ethnic diversity provides an environment enlarging the scope

for individuals to generate new combinations of ideas and resources, thereby increasing the creativity, innovation and productivity of individuals and businesses. In this context, social capital theory helps to explain how this occurs; namely, ethnically diverse neighbourhoods may promote "weak ties", which are an acknowledged source of knowledge exchange, innovation and economic betterment.

The last conflict in BiH (1992–1995) was ethnically characterized and harmful to the ethnic heterogeneity of this society. Still, two decades later, this study reveals positive economic consequences of local ethnic diversity for individuals and households. Consistent with most of the micro-focused literature, we find that ethnic diversity within neighbourhoods – where it exists – is not an economic threat but is rather associated with positive outcomes for individuals and families; in particular, lower probabilities of being in the lowest income categories and higher probabilities of being in medium and higher income categories. These findings have particular implications for poverty reduction. Around 25 per cent of our respondents report no personal income and almost 10 per cent no regular family income in the previous month. Our estimates suggest that living in an ethnically diverse area rather than an ethnically homogeneous area reduces respondents' chances of having no income in the previous month by 4.5 percentage points; while for zero family income in the previous month, the corresponding reduction is around three percentage points. Moreover, at the limit, our estimates imply an income gap of more than 20 per cent between a counter-factual completely heterogenous environment and a counter-factual completely homogenous environment. These estimates are supported by extensive robustness checks, in particular by estimating additional models to allow for heteroskedasticity and to address the potential endogeneity of local ethnic diversity. In each case, the qualitative pattern of estimates remains the same, although quantitatively the effects of ethnic diversity in IV Ordered Probit estimation are much larger. Each estimation strategy has advantages and disadvantages. Hence, with no single dominant strategy, and a strong argument for treating post-conflict ethnic diversity in BiH as exogenous, we prefer the conservative approach of reporting our most modest estimates.

Bosnia and Herzegovina has been a multicultural environment for more than 1,000 years (Malcolm 1996), and our findings suggest that this authentic diversity is economically beneficial for its individuals and families. Conversely, policies and initiatives supporting ethnic homogeneity over diversity – currently present in this long lasting ethnically diverse society – harm the economic performance of individuals and households. A corollary is that policy makers in this post-conflict country, and in similar environments elsewhere – for example, in other Western Balkan countries, including Kosovo, Macedonia, Montenegro and Serbia – should promote ethnic diversity and, across the broad range of public policies, take into consideration the negative effect of ethnic homogeneity. This finding and its corresponding policy implications are consistent with previous research reporting that the business sector in BiH benefits from ethnically diverse surrounding neighbourhoods. When combined, these two findings send a powerful policy message that ethnically more diverse areas in BiH are stronger economically, with positive benefits for individuals, families and businesses. As a consequence, policies designed to encourage and support ethnic inclusion in otherwise homogenous areas have the potential to be welfare enhancing. Hitherto, return policy in BiH mainly failed in its goal of wide-scale return to pre-war homes, leaving instead a substantial part of the country ethnically homogenous. More effort should be made to support return and social inclusion of

ethnic minorities in all parts of the country, developing also an awareness of the economic benefits of diversity.

References

Akerlof, G.A., and Kranton, R. (2010). Identity Economics: How Identities Shape Our Work, Wages, and Well-being. Princeton: Princeton University Press. https://press.princeton.edu/titles/9108.html

Alesina, A., and La Ferrara, E. (2005). Ethnic diversity and economic performance. *Journal of Economic Literature*, 43(3), 762–800. https://www.jstor.org/stable/4129475?seq=1#page_scan_tab_contents

Alesina, A., Baqir, R., and Easterly, W. (1999). Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4), 1243–1284. https://doi.org/10.1162/003355399556269

Barro, R. J., and McCleary, R. (2003). Religion and economic growth. NBER Working Paper 9682. http://www.nber.org/papers/w9682

Bellini, E., Ottaviano, G.I.P., Pinelli, D., and Prarolo, G. (2012). Cultural diversity and economic performance: Evidence from European regions. In R. Crescenzi and M. Percoco (eds.), *Geography, Institutions and Regional Economic Performance*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-33395-8_7

Bieber, F. (2010). Bosnia and Herzegovina since 1990. In Ramet, S.P., *Central and Southeast European Politics Since 1989.* Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511803185.018

Blimes, R.J. (2006). The indirect effect of ethnic heterogeneity on the likelihood of civil war onset. *The Journal of Conflict Resolution*, 50(4), 536–547. https://doi.org/10.1177/00222706289402

Cinnirella, F., Streb, J. (2017). Religious tolerance as engine of innovation. CEPR Discussion Paper No. 12466. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3082292#

Cipolla, C.M. (1972). The diffusion of innovations in early modern Europe. *Comparative Studies in Society and History*, 14(1), 46–52. http://www.jstor.org/stable/178059

Collier, P. (1998). The political economy of ethnicity. CSAE Working Paper Series 1998-08, Centre for the Study of African Economies, University of Oxford. https://ideas.repec.org/p/csa/wpaper/1998-08.html

Constant, A.F., Gataullina, L., and Zimmermann, K.F. (2009). Ethnosizing immigrants. *Journal of Economic Behavior & Organization*, 69(3), 274–287. https://doi.org/10.1016/j.jebo.2008.10.005

Cvitkovic, I. (2017). Nacionalna i konfensionalna slika Bosne i Hercegovine. In I. Cvitkovic, Demografske Promjene u BiH. Sarajevo: Academy of Science and Arts of Bosnia and Herzegovina. http://www.anubih.ba/images/publikacije/posebna_izdanja/ODN/09_posebna_izdanja_CLXXII_9/09_posebna_izdanja_CLXXII_9.pdf

Donia, R.J. (2006). *Sarajevo. A biography*. The University of Michigan Press. https://www.press.umich.edu/189593/sarajevo

Dyrstad, K. (2012). After ethnic civil war - Ethno-nationalism in the Western Balkans. *Journal of Peace Research*, 49(6), 817–831. https://doi.org/10.1177/0022343312439202

Easterly, W. (2001). Can institutions resolve ethnic conflict? *Economic Development and Cultural Change*, 49(4), 687–706. http://www.jstor.org/stable/10.1086/452521

Easterly, W., and Levine, R. (1997). Africa's growth tragedy: policies and ethnic divisions. *Quarterly Journal of Economics*, 112(4), 1203–1250.
https://EconPapers.repec.org/RePEc:oup:qjecon:v:112:y:1997:i:4:p:1203-1250.

Efendic, A. (2010). *Institutions and Economic Performance in Transition Countries - With Special Reference to Bosnia and Herzegovina*. Lambert Academic Publishing.
https://www.researchgate.net/publication/266022605_Institutions_and_economic_performance_in__transition_countries_with_special_reference_to_Bosnia_and_Herzegovina

Efendic, A. (2018). The role of economic and social capital during the floods in Bosnia and Herzegovina. In Dzihic,V. and Solska, M. (eds.) *Crisis Governance in Bosnia and Herzegovina, Croatia and Serbia. The Study of Floods in 2014.* Publisher: Peter Lang.
https://www.researchgate.net/publication/323295545_The_Role_of_Economic_and_Social_Capital_during_the_Floods_in_Bosnia_and_Herzegovina

Efendic, A., Mickiewicz, T., Rebmann, A. (2015). Growth aspirations and social capital: Young firms in a post-conflict environment. *International Small Business Journal*, 33(5), 537–561.
https://doi.org/10.1177/0266242613516987

Efendic, A., Pugh, G., Adnett, N. (2011). Confidence in formal institutions and reliance on informal institutions in Bosnia and Herzegovina - An empirical investigation using survey data. *Economics of Transition*, 19(3), 521–540.
https://EconPapers.repec.org/RePEc:bla:etrans:v:19:y:2011:i:3:p:521-540

Efendic, A. (2016). Emigration intentions in a post-conflict environment - evidence from Bosnia and Herzegovina. *Post-Communist Economies*, 28(3), 335–352.
https://doi.org/10.1080/14631377.2016.1166800

Efendic, A., Babic, B., Rebmann, A. (2017). *Social Capital, Migration, Ethnic Diversity and Economic Performance – Multidisciplinary Evidenced from South-East Europe.* Publisher: Peter Lang
https://www.researchgate.net/publication/318837809_Social_Capital_Migration_Ethnic_Diversity_and_Economic_Performance_-_Multidisciplinary_Evidence_from_South-East_Europe

Escaleras, M., & Register, C. A. (2011). Natural disasters and foreign direct investment. *Land Economics*, 87(2), 346–363. http://www.jstor.org/stable/41307218

Fearon, J. D. (2003). Ethnic and cultural diversity by country. *Journal of Economic Growth*, 8(2), 195–222. http://www.jstor.org/stable/40215943

Federal Ministry of Displaced Persons and Returnees (FMDPR) (2011). Informacija o stanju raseljenih osoba i povratnika u Federaciji Bosne i Hercegovine, iskazanim potrebama i mjerama. Sarajevo: FMDPR. Available at:
https://predstavnickidom-pfbih.gov.ba/upload/file/dokumenti/Informacija_o_stanju_raseljenih_osoba_FBiH.pdf

Foster, E., McCleery, A. (1999). Computer assisted personal interviewing: A method of capturing sensitive information. *IASSIST Quarterly*, 23(2), 26–38.
http://www.websm.org/uploadi/editor/iqvol232forster.pdf

Gertler, M., Florida, R., Gates, G., Vinodrai, T. (2002). Competing on creativity: placing Ontario's cities in North American context. A report prepared for the: Ontario Ministry of Enterprise, Opportunity and Innovation and the Institute for Competitiveness and Prosperity.
http://webarchive.urban.org/publications/410889.html

Gören, E. (2014). How ethnic diversity affects economic growth. *World Development*, 59, July 2014, 275–297. https://doi.org/10.1016/j.worlddev.2014.01.012

Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380. https://doi.org/10.1086/225469

Greene, W.H. (2012, 7ᵗʰ Ed). *Econometric Analysis*. Pearson. https://www.pearson.com/us/higher-education/program/Greene-Econometric-Analysis-7th-Edition/PGM305107.html?tab=overview

Haller, M., Eder, A. (2015). *Ethnic Stratification and Economic Inequality Around the World – the End of Explatation and Esclusion?* Ashgate publishing. https://www.routledge.com/Ethnic-Stratification-and-Economic-Inequality-around-the-World-The-End/collaboration-Eder/p/book/9781409449522

Ibrahim, A. (2017). Social, Human and Job Characteristics as the Determinants of Wages and Gender Disctrimination in Syria: Direct and indirect effects. PhD Thesis, Brunel University London. http://ethos.bl.uk/OrderDetails.do?uin=uk.bl.ethos.715924

Jacobs, J. (1961). *The Death and Life of Great American Cities.* Random House. https://www.goodreads.com/book/show/30833.The_Death_and_Life_of_Great_American_Cities

Korman, S. M. (2015). Examining the effect of religion on economic growth: a partial replication and extension. Honors Theses, 1963–2015. 89. https://digitalcommons.csbsju.edu/honors_theses/89

Lemieux, T. (2006). The "Mincer Equation" thirty years after schooling, experience, and Earnings. In Grossbard, S. (eds.). *Jacob Mincer A Pioneer of Modern Labor Economics.* Springer, Boston, MA. https://doi.org/10.1007/0-387-29175-X_11

Malcolm, N. (1996). *Bosnia – A Short History.* New York University Press https://www.goodreads.com/book/show/1049929.Bosnia

McLeod, P. L., Lobel, S. A., Cox, T. H. (1996). Ethnic diversity and creativity in small groups. *Small Group Research*, 27(2), 248–264. https://doi.org/10.1177/1046496496272003

McLeod, P. L., Lobel, S. A. (1992). The effect of diversity on idea creativity in small groups. *Academy of Management Proceedings*, 1, 227–231. http://proceedings.aom.org/content/1992/1/227.abstract

Mincer, J. A. (1974). *Schooling, Experience and Earnings*. NBER publishing. http://papers.nber.org/books/minc74-1

Mokyr, J. (2016). *A Culture of Growth: The Origins of the Modern Economy.* Princeton University Press. https://press.princeton.edu/titles/10835.html

Montalvo, J. G., Reynal-Querol, M. (2005). Ethnic diversity and economic development. *Journal of Development Economics*, 76(2), 293–323. https://doi.org/10.1016/j.jdeveco.2004.01.002

Norušis, M. J. (2005). *SPSS 13.0 Advanced Statistical Procedures Companion*. Upper Saddle River, NJ: Prentice Hall.

Olzak, S. (2011). Does globalization breed ethnic discontent? *Journal of Conflict Resolution*, 55(1), 3–32. https://doi.org/10.1177/0022002710383666

Osborne, E. (2000). Multiculturalism, and ethnic conflict: A rent-seeking perspective. *Kyklos*, 53(4), 509–526. https://EconPapers.repec.org/RePEc:bla:kyklos:v:53:y:2000:i:4:p:509-25

Ottaviano, G.I.P, Peri, G. (2006). The economic value of cultural diversity: evidence from US cities. *Journal of Economic Geography*, 6(1), 9–44. https://doi.org/10.1093/jeg/lbi002

Patsiurko, N., Campbell, J. L., Hall, J. A. (2012). Measuring cultural diversity: ethnic, linguistic and religious fractionalization in the OECD. *Ethnic and Racial Studies,* 35(2), 195–217. https://doi.org/10.1080/01419870.2011.579136

Posner, D. N. (2000). Measuring ethnic identities and attitudes regarding inter-group relations: Methodological pitfalls and a new technique. A memo prepared for LiCEP2, University of Pennsylvania. http://web.ceu.hu/cps/bluebird/eve/statebuilding/posner.pdf

Rebmann, A., Efendic, A., Mickiewicz, T. (2017). Nascent enterprises and growth aspirations in a post-conflict environment: the role of social capital. In Williams, C. C., Gurtoo, A. (eds.). *Routledge Handbook of Entrepreneurship in Developing Economies*. Routledge Taylor & Francis Group. https://www.researchgate.net/publication/305652095_Nascent_entreprises_and_growth_aspirations_in_a_post-conflict_environment_the_role_of_social_capital

Roodman, D. (2011). Fitting fully observed recursive mixed-process models with cmp. *Stata Journal* 11(2): 159–206. http://www.stata-journal.com/article.html?article=st0224

Sekulic, D., Massey, G., Hodson, R. (1994). Who were the Yugoslavs? Failed sources of a common identity in the former Yugoslavia. *American Sociological Review*, 59(1), 83–97. http://www.jstor.org/stable/2096134

StataCorp (2015). *Stata Base Reference Manual Release 14*. A Stata Press Publication, StataCorp LLC, College Station, Texas. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.637.2432&rep=rep1&type=pdf

Smallbone, D., Kitching, J., Athayde, R. (2010). Ethnic diversity, entrepreneurship and competitiveness in a global city. *International Small Business Journal*, 28(2), 174–190. https://doi.org/10.1177/0266242609355856

UCLA (2015). Stata Data Analysis Examples - Ordered Logistic Regression, UCLA Institute for Digital Research and Development. Available at: http://www.ats.ucla.edu/stat/stata/dae/ologit.htm

UNDP (2000–2010). Early Warning System BiH (quarterly and annual surveys). Sarajevo: UNDP BiH.

Vanhoutte, B., Hooghe, M. (2012). Do diverse geographical contexts lead to diverse friendship networks? A multilevel analysis of Belgian survey data. *International Journal of Intercultural Relations*, 36(3), 343-352. https://doi.org/10.1016/j.ijintrel.2011.09.003

Williams, R. A. (2006). Generalized Ordered Logit/Partial Proportional Odds Models for Ordinal Dependent Variables. *The Stata Journal* 6(1), 58–82. http://www.stata-journal.com/abstracts/st0097.pdf

Williams, R. A. (2010). Fitting heterogeneous choice models with oglm. *The Stata Journal*, 10(4) 540–567. http://www.stata-journal.com/article.html?article=st0208

Wooldridge, J. M. (2010). *Econometric Analysis of Cross Section and Panel Data*. MIT Press. https://mitpress.mit.edu/books/econometric-analysis-cross-section-and-panel-data

Appendix 1: Survey questionnaire

INDICATORS	Formulation of the questions used in survey
Total monthly income of the respondents (the same scale is used for personal income and family income) BAM is BiH currency with the exchange rate 1 Euro=1.95 BAM	Denoted in BAM, what was your total income last month including all sources of income: 1. No income in the previous month 2. Up to 100 3. From 101 to 386 4. From 387 to 500 5. From 501 to 700 6. From 701 to 900 7. From 901 to 1100 8. From 1101 to 1500 9. From 1501 to 2000 10. From 2001 to 3000 11. Over 3000 12. Do not know
Total monthly income of the family	Denoted in BAM, what was the total income of your family last month including all sources of income: (the same scale)
Ethnic diversity	How would you describe the level of ethnic diversity in your neighbourhood: a. There is only one ethnicity b. Small number of people from other ethnicities c. Larger number of other ethnicities d. There is a large diversity (i.e., it is very mixed) e. Do not know
Ethnic diversity of networks	What percentage of your friends belongs to an ethnic group other than your own;_____%
Minority v.v. Majority ethnic status	In the area where you live (street or village) is your ethnic group with which you identify yourself the minority or majority: a. Minority b. Majority c. Do not know
War migration status	During the war in B&H, did you ever leave your home? a. Yes, but we came back rapidly b. Yes, but we came back few years after the war c. Yes, and we never came back d. No, we did not leave our home during the war e. Do not know
Experience of being abroad	Over the last 3 years, how much time did you spend outside of countries from Former Yugoslavia? a. No time at all b. Few days c. Up to a week d. Up to a month e. Up to a year f. More than a year g. Do not know
Age	How old are you?
Gender	What is your gender? A. Male b. Female
Marital status	What is your marital status? a. Single b. Married c. Cohabitation d. Separated e. Widowed f. Do not know
Education	What is your highest level of education? a. No education or incomplete elementary education b. Elementary education c. High school/Gymnasium d. College e. University education f. Master/Doctoral degree g. Do not know
Area	urban – city, suburban and rural
Municipality	Municipality: _____

Appendix 2: Frequency and calculation of the average wage effects

Personal income OP model

EARNING SCALE	FREQUENCY FROM THE REGRESSION	MID INCOME	ESTIMATED PROBABILITIES FOR EACH CATEGORY, d_diversity=0	ESTIMATED PROBABILITIES FOR EACH CATEGORY, d_diversity=1	AVERAGE INCOME d_diversity=0	AVERAGE INCOME d_diversity=1
1	461	0	0.316	0.204	0.0	0.0
2	57	50	0.034	0.027	1.7	1.4
3	534	244	0.301	0.284	73.3	69.3
4	304	444	0.153	0.181	68.0	80.3
5	192	600	0.088	0.120	52.6	72.0
6	109	800	0.045	0.069	36.0	55.0
7	79	1000	0.030	0.051	30.5	50.9
			0.967	**0.937**	**262.2**	**329.0**
Effect of diversity on mean income if BiH changes from counter-factual completely homogenous to counter-factual completely heterogenous environment (first seven of the 11 income categories, capturing around 95% of observations)=25.5%						

Personal income heteroskedasticity OP model

EARNING SCALE	FREQUENCY FROM THE REGRESSION	MID INCOME	ESTIMATED PROBABILITIES FOR EACH CATEGORY, d_diversity=0	ESTIMATED PROBABILITIES FOR EACH CATEGORY, d_diversity=1	AVERAGE INCOME d_diversity=0	AVERAGE INCOME d_diversity=1
1	461	0	0.315	0.197	0.0	0.0
2	57	50	0.037	0.029	1.8	1.4
3	534	244	0.307	0.302	74.9	73.8
4	304	444	0.144	0.188	63.8	83.6
5	192	600	0.082	0.118	49.2	70.8
6	109	800	0.043	0.064	34.7	51.3
7	79	1000	0.031	0.046	31.0	46.0
			0.959	**0.944**	**255.4**	**327.0**
Effect of diversity on mean income if BiH changes from counter-factual completely homogenous to counter-factual completely heterogenous environment (first seven of the 11 income categories, capturing around 95% of observations)=21.9%						

Economics
The Open-Access, Open-Assessment E-Journal

Please note:

You are most sincerely encouraged to participate in the open assessment of this article. You can do so by either recommending the article or by posting your comments.

Please go to:

The Editor